Advanced Engineering Informatics 52 (2022) 101560

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Full length article

SRS: Spatial-tagged radio-mapping system combining LiDAR and mobile-phone data for indoor location-based services

Yu-Cheol Lee *

Artificial Intelligence Laboratory, ETRI, Daejeon 34129, Republic of Korea
Artificial Intelligence, University of Science and Technology, Daejeon 34113, Republic of Korea
Department of Computer Science, Stony Brook University, New York 11794, USA

ARTICLE INFO

Keywords:
RSS fingerprint
Radio map
Map building
Fingerprinting localization
Location-based service

ABSTRACT

One of the most challenging issues in radio received signal strength (RSS)-based localization systems is the generation and distribution of a radio map with a coordinate system linked with spatial information in a large indoor space. This study proposes a novel spatial-tagged radio-mapping system (SRS) that effectively combines the heterogeneous properties of LiDAR and mobile phones to simultaneously perform both spatial and radio mappings. The SRS consists of synchronization, localization, and map building processes, and enables real-time spatial and radio mapping. In the synchronization process, the distance range, motion data, and radio signals obtained through the LiDAR and mobile phone are collected in nodal units according to the sensing time. In the localization process, a feature variance filter is used to control the number of features generated from LiDAR and estimate the positions at which the nodes are generated in real time according to the motion data and radio signals. In map building, the estimated positions of the nodes are used to extract spatial and radio maps by using a unified location coordinate system. To ensure mobility, the SRS is manufactured in the form of a backpack supporting LiDAR and a mobile phone; the usefulness of the system is experimentally verified. The experiments are performed in a large indoor shopping mall with a complex structure. The experimental results demonstrated that a common coordinate system could be used to build spatial and radio maps with high accuracy and efficiency in real time. In addition, the field applicability of the SRS to location-based services is experimentally verified by applying the constructed radio map to well-known fingerprinting algorithms using the heterogeneous mobile phones.

1. Introduction

The scope of the localization technology has recently been expanded to large indoor spaces [1]. The increase in the size and complexity of buildings has resulted in the increasing demand of localization technologies to provide information for destinations indoors. The indoor localization technology is already widely used in real life; representative applications include customer information services in shopping centers, surveillance and reconnaissance services in security facilities, and delivery services in distribution warehouses [2]. In particular, with increasing use of mobile phones and developments in telecommunication devices, indoor localization technology is being closely utilized in the field of life safety, e.g., for emergency calls and emergency exit guidance [3].

With the proliferation of pervasive network devices, the fingerprinting method based on the received signal strength (RSS) of WiFi and Bluetooth has become the most widely used indoor localization technology [4,5]. Other methods include measurement of the time difference of arrival [6] and angle of arrival [7]. However, their widespread use in the public domain is limited owing to the high cost of purchasing, installing, managing, and maintaining specially designed devices such as ultra-wideband (UWB) [8,9]. Conversely, the fingerprinting method can implement the RSS of WiFi and Bluetooth devices without the need of additional devices, which are representative pervasive network devices installed for cloud computing. Hence, the RSS-based fingerprinting method, which does not require additional costs even in the case of mobile phones, is being developed into a practical indoor localization technology.

Three major technical considerations are required to widely apply the radio RSS-based fingerprinting localization method: (1) fingerprint accuracy of the radio map containing the RSS information [10], (2) connectivity with the spatial map [11], and (3) map building efficiency [12]. Inadequacies in any one of these factors complicates the application of fingerprint localization to the field. Therefore, a solution is required that simultaneously addresses all three considerations.

* Correspondence to: Artificial Intelligence Laboratory, ETRI, Daejeon 34129, Republic of Korea
E-mail addresses: yclee@etri.re.kr, yu-cheol.lee@stonybrook.edu.

https://doi.org/10.1016/j.aei.2022.101560
Received 20 August 2021; Received in revised form 6 February 2022; Accepted 14 February 2022
Available online 26 February 2022
1474-0346/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Regarding the first consideration, the most important factor in the radio RSS-based localization method is to build an accurate radio map. This is because given the principles of localization in the fingerprinting method, the currently received radio signal must be matched with the RSS stored in the radio map. Therefore, studies have developed various techniques for building radio maps [13,14].

Second, the location of the fingerprint in the radio map must be registered to the same coordinate system as that in the spatial map [11]. The fingerprinting localization result is typically output as the location at which the RSS is collected. In practical applications, it is necessary to determine where the location estimated using the radio map corresponds in space. The method's utility will inevitably decrease if the localization result obtained through the radio map is not linked with the location on the spatial map used as the user interface [15].

Finally, in terms of cost, the building technique for the radio map must have a large mapping space size per hour. The time required to build the radio map is directly related to the cost of implementing localization [12,16]. Owing to the growing size of indoor spaces, the efficiency of building radio maps is emerging as a crucial technical factor in the proliferation of location-based services (LBSs).

As a solution to these three issues, this study proposes a novel spatial-tagged radio-mapping system (SRS) that can simultaneously build spatial and radio maps in real time by fusing heterogeneous sensing data obtained from LiDAR and mobile phone. Essentially, SRS synchronizes and collects the distance range and radio signals acquired from LiDAR and mobile phone as one node. SRS uses the LiDAR distance range to estimate the collected node position and can use this position to simultaneously build the spatial and radio maps. This system is technically significant in that it combines the advantages of LiDAR and mobile phone to provide a practical mapping tool for effectively building and distributing the radio map for indoor localization.

Furthermore, SRS is considered an appropriate solution for the three technical contributions owing to its improved mapping efficiency, reliable accuracy, and connectivity with the spatial map when building the radio map. Considering that the spatial and radio maps are built using the position estimated through LiDAR-based localization technology, the location coordinate system of both maps is automatically registered to a common coordinate system. Hence, the fingerprint location of the radio map can ensure reliable accuracy. This study also developed a feature variance filter (FVF) that controls the number of features in the localization module by identifying trends of a mobile phone's motion data and radio signals. FVF is a core technology for achieving efficient and real-time mapping, and it enables mapping at even walking speeds using SRS.

The remainder of this paper is organized as follows. After a discussion of the related work in the following section, Section 3 presents the detailed algorithm of SRS. Section 4 describes experiments and results to verify the major contributions of SRS. Section 5 concludes the research work and outlines possible future studies on practical applications of SRS.

2. Related works

To implement radio RSS-based indoor localization technology, a radio map whose coordinate system is linked with the spatial map must be provided beforehand. This is typically implemented by designating fingerprint location containing the collected radio received signal information on the floorplan used as the spatial map [17,18]. A floorplan is used because the direct construction of a spatial map containing geometric data using only equipment, such as mobile phone, is difficult. Representative approaches include point manual calibration (PMC) [19,20], simultaneous localization and mapping (SLAM) [21], learning-based approach [22,23], and cloud collection [24]. Actual localization is possible only with a radio map; however, as LBS can only be used when the location can be defined in space, registration using the spatial map is an important technical factor.

The PMC estimates the fingerprint location using pedestrian dead-reckoning (PDR) that combines inertial measurement unit (IMU), gyroscope, and magnetometer, and links with the map coordinate system. Usually, a straight line consisting of departure and arrival points is manually designated on the map, and the operator estimates the position of the points constituting the actual straight path with a constant stride by PDR. By manual scale adjustment and calibration of the location of the estimated points on the spatial map, fingerprint location information with collected radio signals can be linked with the spatial map. Because the operator manually performs matching with the spatial map, the method is highly accurate, and is thus used by many researchers for constructing the reference dataset of fingerprinting localization performance. However, the involvement of manual work leads to low efficiency [25–27].

While the SLAM technique is highly efficient, the registration accuracy between the floorplan and radio map is low. After designating the starting point on the drawing, this method connects the movement trajectory using radio signals and the PDR and expands the mapping area. This is more efficient compared to the PMC method because there are fewer manual connection tasks. However, large errors arise in the location estimated using PDR in the case of a non-straight movement; this results in the problem of incorrect loop closure in SLAM [28,29]. As a result, the floorplan and actual radio fingerprint location are frequently inconsistent. In some investigations, to address the limitations of PDR, SLAM using a LiDAR and a panoramic camera is employed to construct a spatial map, and then it is manually linked with a radio map [30]. The spatial map is used only for extracting the ground-truth location information of the waypoints for the challenge. The operator constructs a radio map for the location of the waypoints defined on the spatial map with PMC method. LiDAR-based SLAM is applied to the construction of radio maps, but a system that can build both spatial and radio maps concurrently as in the case of SRS is not provided.

A learning-based approach designates some labeled fingerprint locations on the spatial map and adds fingerprint information for unknown areas according to the learning model. The method enables building a radio map of the entire area by specifying a few labeled fingerprints, thus reducing manual calibration cost compared to PMC. However, the method requires efforts to designate the labeled fingerprint locations on the spatial map. Usually, this method can be used only in indoor environments that include a place where the reference location can be known, such as GPS. [20].

Cloud collection is a method for creating and updating the radio fingerprint location in space through a complex technique, such as with PMC, SLAM, and learning-based approach [31]. Even in large spaces, the mapping efficiency increases with increase in the number of mapping participants, and the previously collected fingerprint information can be continuously updated. However, ensuring consistent accuracy is difficult because the radio map is built using various techniques.

In the four aforementioned techniques, majority of the methods for registering the spatial and radio maps require a floorplan to be provided beforehand. Even if a floorplan is given, the radio fingerprint collection location is only registered in the floorplan. To improve the efficiency of fingerprint creation, an interpolation method that considers the physical distance of the floorplan has been used [32]. Moreover, to increase the fingerprint accuracy, the radio signals are collected to reflect the structural shapes of the floorplan [33]. However, these methods are inapplicable to situations in which a spatial map, such as a floorplan, is not provided.

Accordingly, to address the limitations of the existing techniques, in this study, we propose SRS, which combines the data acquired from LiDAR and mobile phone to simultaneously build and distribute the spatial and radio maps. Additionally, to confirm the applicability of SRS as a base technology for indoor LBS, its performance was experimentally verified in terms of securing radio-map fingerprint accuracy, registration of location information using the spatial map, and improvement of mapping efficiency.

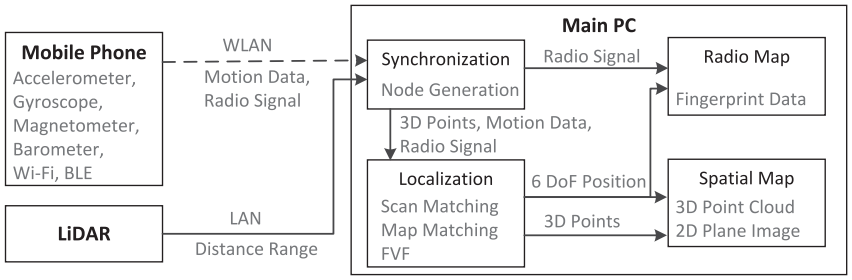


Fig. 1. Overview of the proposed spatial-tagged radio-mapping system.

3. Spatial-tagged radio-mapping system

SRS can simultaneously build spatial and radio maps in an integrated coordinate system by fusing the data measured from LiDAR and mobile phone. Fig. 1 describes the components and operation process of the SRS. In the main PC, the synchronization, localization, and map building modules are executed step by step.

In the synchronization step, the radio signals and the distance range acquired from mobile phones and LiDAR connected through WLAN and LAN, respectively, are synchronized according to the acquisition time and stored in one node. In the localization step, the 3D points converted from the LiDAR distance range are used to estimate the position at six degrees of freedom (DoF) through scan and map matchings. In the mapping step, the estimated 6-DoF position and 3D points are used to construct a 3D-point cloud and 2D-plane image as the spatial map. The estimated 6-DoF position and radio signals can be used to build the fingerprint information as a radio map. This study also proposes an FVF that can analyze whether the moving space changes according to the motion data and radio signals measured from mobile phones in the localization module. Using FVF to control the number of features to be employed for localization matching, the map can be built in real time.

3.1. Synchronization

LiDAR and mobile phone are physically different devices connected via LAN and WLAN, respectively. As such, synchronization is vital to enhance the accuracy of the collected measurement information. In particular, as mobile phone is connected via WLAN, the collected data are likely to be delayed. Therefore, to synchronize the sensing data, the 3D points converted from the 3D range distances of LiDAR, $\mathbf{P}_t^L$, and the motion data and radio signals $\mathbf{Z}_t$ from mobile phone are composed as one node, $\mathbf{N}_t$, according to the measurement time as follows:

$$\mathbf{N}_t = \left\{\mathbf{P}_t^L, \mathbf{Z}_t\right\}. \tag{1}$$

In addition, the information obtained from the mobile phone, $\mathbf{Z}_t$, can be transmitted to the main PC through WLAN, which may cause latency. To overcome the latency issue, nodes are created based on the time the data acquired from the device and not the time data sent to the main PC. Specifically, in LiDAR and mobile phones, the time of data acquisition is passed to the main PC and saved in a buffer, and these are constructed as a single node when the acquisition time of $\mathbf{P}_t^L$ and $\mathbf{Z}_t$ is within a set range.

3D points data $\mathbf{P}_t^L$ is the local coordinate system with the LiDAR position as the origin, with N_a channels in the vertical direction and N_p points, $(x_{n,i}^L, y_{n,i}^L, z_{n,i}^L)$, for each channel in the horizontal direction. This system is represented as

$$\begin{aligned}\mathbf{P}_t^L &= \left\{\mathbf{p}_{n,i}^L \middle| n = 1,\dots,N_a, i = 1,\dots,N_p\right\},\\ \mathbf{p}_{n,i}^L &= \left(x_{n,i}^L, y_{n,i}^L, z_{n,i}^L\right).\end{aligned} \tag{2}$$

Motion data $\mathbf{m}_t$ of $\mathbf{Z}_t$ consists of the sum of the accelerations of the three-axis, a_t; sum of the angular velocities of the three-axis, ω_t; sum of the magnetic fields of the three-axis, μ_t; and atmospheric pressure ρ_t.

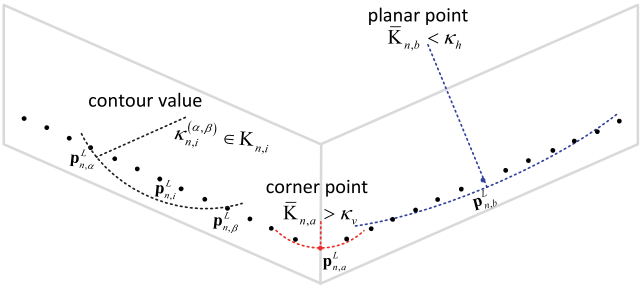


Fig. 2. Point feature extraction based on the sequential contour values.

In addition, radio signal $\mathbf{r}_t$ of $\mathbf{Z}_t$ consists of N_z MAC addresses and RSS values of (mc_j, r_j) acquired from nearby Wi-Fi and Bluetooth devices:

$$\begin{aligned}\mathbf{Z}_t &= \left(\mathbf{m}_t, \mathbf{r}_t\right), \mathbf{m}_t = \left(a_t, \omega_t, \mu_t, \rho_t\right),\\ \mathbf{r}_t &= \left\{\left(mc_j, r_j\right) \middle| j = 1,\dots,N_z\right\}.\end{aligned} \tag{3}$$

3.2. Localization for map construction

3.2.1. Feature extraction

We consider that only the features of major static objects should be extracted from the LiDAR-based 3D points and used for matching. This is because the use of all the point data of LiDAR for localization matching might not guarantee the real-time performance owing to heavy computational burden. Moreover, for objects that move dynamically, such as humans, frequent changes in location may adversely impact the matching accuracy. Therefore, only the point features of corners and planes extracted from static objects, such as columns and walls, with fixed locations are used for matching.

The LiDAR technology can measure the locations of nearby objects in the form of points at uniform bearing angle intervals. Thus, using the curvature value of sequential points, the point features corresponding to the corners and planes can be extracted, as shown in Fig. 2. The curvature value for given point $\mathbf{p}_{n,i}^L$ in the nth scan channel can be calculated using scan points $\mathbf{p}_{n,\alpha}^L$ and $\mathbf{p}_{n,\beta}^L$ on either side as follows:

$$\kappa_{n,i}^{(\alpha,\beta)} = \frac{\left\|\left(\mathbf{p}_{n,\alpha}^L, \mathbf{p}_{n,\beta}^L\right)\right\|}{\left\|\mathbf{p}_{n,i}^L\right\|}. \tag{4}$$

To determine whether this corresponds to the point features of $\mathbf{p}_{n,i}^L$, curvature set $\mathrm{K}_{n,i}$ is calculated as follows:

$$\mathrm{K}_{n,i} = \left\{\kappa_{n,i}^{(\alpha,\beta)} \middle| \frac{(\alpha+\beta)}{2} = i, \cos^{-1}\left(\frac{\mathbf{p}_{n,\alpha}^L \cdot \mathbf{p}_{n,\beta}^L}{\left\|\mathbf{p}_{n,\alpha}^L\right\|\left\|\mathbf{p}_{n,\beta}^L\right\|}\right) < \frac{\pi}{4}\right\}, \tag{5}$$

where α and β form an index pair, i, in the nth channel. Given that majority of corner shapes are vertical, the relative angle between $\mathbf{p}_{n,\alpha}^L$ and $\mathbf{p}_{n,\beta}^L$ spans up to a maximum of $\pi/4$. If the average value of curvature set $\bar{\mathrm{K}}_{n,i}$ is larger than that of corner threshold κ_v or smaller

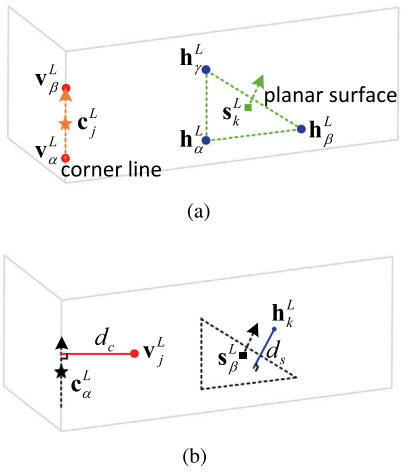


Fig. 3. (a) Geometrical feature extraction. (b) Extraction of the matching distance between the current point and geometrical features of different scans.

than that of plane threshold κ_h, $\mathbf{p}_{n,i}^L$ can be defined as a corner point $\mathbf{v}_j^L$ or a planar point $\mathbf{h}_k^L$ as follows:

$$\mathbf{V}_t^L = \left\{ \mathbf{v}_j^L \,|j = 1, \ldots, N_v \right\} = \left\{ \mathbf{p}_{n,i}^L \,|\bar{\mathrm{K}}_{n,i} > \kappa_v, \mathbf{p}_{n,i}^L \in \mathbf{P}_t^L \right\}, \tag{6}$$

$$\mathbf{H}_t^L = \left\{ \mathbf{h}_k^L \,|k = 1, \ldots, N_h \right\} = \left\{ \mathbf{p}_{n,i}^L \,|\bar{\mathrm{K}}_{n,i} < \kappa_h \,, \mathbf{p}_{n,i}^L \in \mathbf{P}_t^L \right\}, \tag{7}$$

where $\mathbf{V}_t^L$ and $\mathbf{H}_t^L$ are the point feature groups comprising N_v corner points and N_h planar points, respectively.

The corner lines and planar surfaces can be extracted as the geometrical features of the space using the points corresponding to the corners and planes. Fig. 3(a) presents the extraction of the corner lines and planar surfaces in detail. The center point of corner line $\bar{\mathbf{v}}_j^L$ and unit vector $\hat{\mathbf{u}}_j^L$ can be calculated using two corner points $\mathbf{v}_\alpha^L$ and $\mathbf{v}_\beta^L$ nearest to each other as:

$$\mathbf{C}_t^L = \left\{ \mathbf{c}_j^L \,|j = 1, \ldots, N_c \right\} = \left\{ \left(\bar{\mathbf{v}}_j^L, \hat{\mathbf{u}}_j^L \right) \middle| \bar{\mathbf{v}}_j^L = \frac{\mathbf{v}_\alpha^L + \mathbf{v}_\beta^L}{2}, \hat{\mathbf{u}}_j^L = \frac{\overrightarrow{\mathbf{v}_\alpha^L \mathbf{v}_\beta^L}}{\left| \overrightarrow{\mathbf{v}_\alpha^L \mathbf{v}_\beta^L} \right|} \right\}, \tag{8}$$

where $(\mathbf{v}_\alpha^L, \mathbf{v}_\beta^L) \in \mathbf{V}_t^L$, and α and β are different indices. In addition, N_c sets of $\mathbf{c}_j^L$ constitute the corner line group, $\mathbf{C}_t^L$. Further, the center point of planar surface $\bar{\mathbf{h}}_k^L$ and normal unit vector $\hat{\mathbf{n}}_k^L$ can be calculated using the three planar points, $\mathbf{h}_\alpha^L$, $\mathbf{h}_\beta^L$, and $\mathbf{h}_\gamma^L$, that are nearest to each other as follows:

$$\begin{aligned} \mathbf{S}_t^L &= \left\{ \mathbf{s}_k^L \,|k = 1, \ldots, N_s \right\} \\ &= \left\{ \left(\bar{\mathbf{h}}_k^L, \hat{\mathbf{n}}_k^L \right) \middle| \bar{\mathbf{h}}_k^L = \frac{\mathbf{h}_\alpha^L + \mathbf{h}_\beta^L + \mathbf{h}_\gamma^L}{3}, \hat{\mathbf{n}}_k^L = \frac{\overrightarrow{\mathbf{h}_\alpha^L \mathbf{h}_\beta^L} \times \overrightarrow{\mathbf{h}_\alpha^L \mathbf{h}_\gamma^L}}{\left| \overrightarrow{\mathbf{h}_\alpha^L \mathbf{h}_\beta^L} \times \overrightarrow{\mathbf{h}_\alpha^L \mathbf{h}_\gamma^L} \right|} \right\}, \end{aligned} \tag{9}$$

where $(\mathbf{h}_\alpha^L, \mathbf{h}_\beta^L, \mathbf{h}_\gamma^L) \in \mathbf{H}_t^L$, and α, β, and γ are different indices. In addition, N_s number of sets of $\mathbf{s}_k^L$ constitute the planar surface group, $\mathbf{S}_t^L$.

3.2.2. Local position estimation

Scan matching based on sequential LiDAR-based 3D points estimates the 6-DoF position in the local coordinate system as follows:

$$\mathbf{x}_t^L = \begin{bmatrix} x_t^L & y_t^L & z_t^L & \phi_t^L & \theta_t^L & \varphi_t^L \end{bmatrix}^T, \tag{10}$$

where (x_t^L, y_t^L, z_t^L) signifies the position in the 3D space and $(\phi_t^L, \theta_t^L, \varphi_t^L)$ indicates the roll, pitch, and yaw angles, respectively. Assuming that LiDAR moves from $\mathbf{x}_{t-1}^L$ up to $\mathbf{T}$ and rotates up to $\mathbf{R}$, the LiDAR position can be defined as $\mathbf{x}_t^L$, and is calculated as

$$\begin{aligned} \mathbf{x}_t^L(4:6) &= \mathbf{x}_{t-1}^L(4:6) + \mathbf{R}, \\ \mathbf{x}_t^L(1:3) &= \hat{\mathbf{R}} \left[\mathbf{x}_{t-1}^L + \mathbf{T} \right], \end{aligned} \tag{11}$$

where $\mathbf{T}$ and $\mathbf{R}$ express translation distance $[t_x, t_y, t_z]^T$ and rotation angle $[r_x, r_y, r_z]^T$, respectively. $\hat{\mathbf{R}}$ is the skew-symmetric matrix calculated according to Rodrigues rotation matrix formula using rotation angle $\mathbf{R}$ [34]. To calculate $\mathbf{T}$ and $\mathbf{R}$, the matching distance between the previous geometrical feature group, $(\mathbf{C}_{t-1}^L, \mathbf{S}_{t-1}^L)$, and the current point feature group, $(\mathbf{V}_t^L, \mathbf{H}_t^L)$, can be calculated as shown in Fig. 3(b). For $\mathbf{c}_\alpha^L$ and $\mathbf{s}_\beta^L$, $(\bar{\mathbf{v}}_\alpha^L, \hat{\mathbf{u}}_\alpha^L)$ and $(\mathbf{h}_\beta^L, \hat{\mathbf{n}}_\beta^L)$ are used as elements, respectively:

$$d_c = f(\mathbf{C}_{t-1}^L, \mathbf{V}_t^L) = \left| \overrightarrow{\bar{\mathbf{v}}_\alpha^L \mathbf{v}_j^L} \times \hat{\mathbf{u}}_\alpha^L \right|, \tag{12}$$

$$d_s = f(\mathbf{S}_{t-1}^L, \mathbf{H}_t^L) = \left| \overrightarrow{\bar{\mathbf{h}}_\beta^L \mathbf{h}_k^L} \cdot \hat{\mathbf{n}}_\beta^L \right|, \tag{13}$$

where the nearest corner line, $\mathbf{c}_\alpha^L \in \mathbf{C}_{t-1}^L$, and planar surface $\mathbf{s}_\beta^L \in \mathbf{S}_{t-1}^L$ are from corner point $\mathbf{v}_j^L \in \mathbf{V}_t^L$ and planar point $\mathbf{h}_k^L \in \mathbf{H}_t^L$, respectively.

To calculate $\mathbf{R}$ and $\mathbf{T}$ with the smallest value of $\mathbf{d}_t$ composed of d_c and d_s, the moving position of LiDAR, $\mathbf{x}_t^L$, in the local coordinate system can be estimated as follows:

$$\mathbf{x}_t^L = \underset{[\mathbf{R}|\mathbf{T}]}{\arg\min} \left(\mathbf{d}_t \right). \tag{14}$$

$\mathbf{R}$ and $\mathbf{T}$ can be defined as a nonlinear least-squares problems due to their dependent relationship and are calculated using the Levenberg–Marquardt algorithm, which combines the Gauss–Newton and gradient-descent methods and finds a solution relatively quickly [35,36].

$$\mathbf{Q}_t \leftarrow \mathbf{Q}_t + \left(\mathbf{J}_t^T \mathbf{J}_t + \mu \mathbf{I} \right)^{-1} \mathbf{J}_t^T \mathbf{d}_t, \tag{15}$$

where μ is the non-negative damping factor. If $\mathbf{Q}_t = [\mathbf{R}|\mathbf{T}]$, then $\mathbf{J}_t$ signifies the Jacobian matrix of $f_c(\cdot)$ and $f_s(\cdot)$ for six elements $[t_x, t_y, t_z, r_x, r_y, r_z]^T$ of $\mathbf{Q}_t$. In addition as $\mathbf{v}_j^L$ and $\mathbf{h}_k^L$, which are the variables of $f_c(\cdot)$ and $f_s(\cdot)$, the Jacobian matrix is developed according to the chain rule as follows:

$$\mathbf{J}_c = \frac{\partial f_c}{\partial \mathbf{Q}_t} = \frac{\partial f_c}{\partial \mathbf{v}_j^L} \frac{\partial \mathbf{v}_j^L}{\partial \mathbf{Q}_t}, \; \mathbf{J}_s = \frac{\partial f_s}{\partial \mathbf{Q}_t} = \frac{\partial f_s}{\partial \mathbf{h}_k^L} \frac{\partial \mathbf{h}_k^L}{\partial \mathbf{Q}_t}. \tag{16}$$

$\mathbf{J}_t$ consists of $\mathbf{J}_c$ and $\mathbf{J}_s$ according to the type of point feature. $\mathbf{Q}_t$ can be calculated using (15) and $\mathbf{x}_t^L$ can be estimated using (11).

3.2.3. Global position estimation

Position $\mathbf{x}_t^L$ estimated through scan matching is close to the dead-reckoning data, which accumulates error. In fact, the scan matching enables a high possibility for accumulating errors because it only uses the features of the previous step sequential for local position estimation [37]. For instance, if there are few features locally or many features extracted from moving objects, the position errors of scan matching might occur owing to the use of the local limited features.

In contrast, map matching can effectively solve the error accumulation problem in scan matching by using features extracted from various times and locations stored as the map in the world coordinate system. Through map matching using the features in the world coordinate system, the global position without the error accumulation should be estimated as follows:

$$\mathbf{x}_t^W = \begin{bmatrix} x_t^W & y_t^W & z_t^W & \phi_t^W & \theta_t^W & \varphi_t^W \end{bmatrix}^T. \tag{17}$$

For the efficiency of map matching calculation, the global position can be predicted using relative position $\Delta\mathbf{x}_t$ between times $t-1$ and t in the local coordinate system as follows:

$$\bar{\mathbf{x}}_t^W = \mathbf{x}_{t-1}^W + \Delta\mathbf{x}_t, \; \Delta\mathbf{x}_t = \mathbf{x}_t^L - \mathbf{x}_{t-1}^L. \tag{18}$$

Using predicted global position $\tilde{\mathbf{x}}_t^W$, feature points $(\mathbf{V}_t^L, \mathbf{H}_t^L)$ in the local coordinate system can be transformed into the world coordinate system as follows:

$$\mathbf{V}_t^W = \left\{ \mathbf{v}_j^W \,\middle|\, \mathbf{v}_j^W = \mathbf{R}(\tilde{\mathbf{x}}_t^W)\left[\mathbf{v}_j^L + \mathbf{T}(\tilde{\mathbf{x}}_t^W)\right] \right\}, \tag{19}$$

$$\mathbf{H}_t^W = \left\{ \mathbf{h}_k^W \,\middle|\, \mathbf{h}_k^W = \mathbf{R}(\tilde{\mathbf{x}}_t^W)\left[\mathbf{h}_k^L + \mathbf{T}(\tilde{\mathbf{x}}_t^W)\right] \right\}, \tag{20}$$

where translation $\mathbf{T}(\tilde{\mathbf{x}}_t^W)$ and rotation $\mathbf{R}(\tilde{\mathbf{x}}_t^W)$ correspond to the first to third elements $\tilde{\mathbf{x}}_t^W(1{:}3)$ and fourth to sixth elements $\tilde{\mathbf{x}}_t^W(4{:}6)$ of $\tilde{\mathbf{x}}_t^W$, respectively. In addition, if the geometrical feature map, comprising geometrical features of the world coordinate system up to time t', is defined as $(\mathbf{C}_{t'}^M, \mathbf{S}_{t'}^M)$, then the global position $\mathbf{x}_t^W$ can be estimated by matching the map with the current point feature group, $(\mathbf{V}_t^W, \mathbf{H}_t^W)$. This map matching follows the same process from (11) to (16) in Section 3.2.2.

Unlike scan matching, map matching utilizes the geometrical maps accumulated up to a certain time. Both scan and map matchings require a step in which the nearest geometrical feature is determined from the point feature. Unlike scan matching, the map matching uses a map with accumulated geometrical features and hence, searching requires a considerable amount of computation even when using an algorithm such as a k-d tree [38]. As such, real-time mapping is a challenging task.

Accordingly, in this study, an FVF is proposed that effectively controls the number of features in the geometrical map by detecting motion and spatial changes according to the variations in the sequential motion data and radio signals of $\mathbf{Z}_t$ and $\mathbf{Z}_{t-1}$ acquired from mobile phone. The probability that FVF λ_t comprises motion probability $p_m(\Delta\mathbf{m}_t)$ and spatial probability $p_r(\Delta\mathbf{r}_t)$ can be confirmed with only slight computational complexity as follows:

$$\lambda_t = p(\mathbf{Z}_t, \mathbf{Z}_{t-1}) = p_m(\Delta\mathbf{m}_t) \cdot p_r(\Delta\mathbf{r}_t), \tag{21}$$

where the motion probability is expressed as the product of the probabilities according to the changes in acceleration Δa_t, angular velocity $\Delta\omega_t$, magnetic field $\Delta\mu_t$, and ambient atmospheric pressure $\Delta\rho_t$ between $\mathbf{Z}_{t-1}$ and $\mathbf{Z}_t$:

$$\begin{aligned} p_m(\Delta\mathbf{m}_t) &= \bar{p}_a(\Delta a_t) \cdot \bar{p}_\omega(\Delta\omega_t) \cdot \bar{p}_\mu(\Delta\mu_t) \cdot \bar{p}_\rho(\Delta\rho_t), \\ \bar{p}_{[\cdot]}(\Delta) &= \left(1 - \frac{1}{\sqrt{2\pi\sigma_{[\cdot]}^2}} \exp\frac{-(\Delta)^2}{2\sigma_{[\cdot]}^2}\right), \end{aligned} \tag{22}$$

where $\bar{p}_{[\cdot]}$ indicates the normal probability density function with variance $\sigma_{[\cdot]}^2$. In addition, spatial probability $p_r(\Delta\mathbf{r}_t)$ is calculated according to the difference in the RSS value as follows:

$$\begin{aligned} p_r(\Delta\mathbf{r}_t) &= \bar{p}_r\left(\sum(\Delta r_{jk})\right), \\ \Delta r_{jk} = \left|r_j - r_k\right| &= \left|\left(10^{27.55-20\log(f)}\right)\left(10^{|r_j|-|r_k|}\right)\right|, \end{aligned} \tag{23}$$

where $(mc_j, r_j) \in \mathbf{Z}_t$ and $(mc_k, r_k) \in \mathbf{Z}_{t-1}$, and RSS values r_j and r_k in which $mc_j = mc_k$ are selected. Here, residual value is calculated using a formula that transforms the RSS from decibel units to metric units based on the free-space path loss [39]. Metric units can be used to accurately detect the physical changes in space. The RSS with low accuracy is used rather than distances obtained from LiDAR because the radio signals are closely affected by spatial changes. i.e., it is possible to effectively detect whether there are changes in the movement space through changes in distances of RSS.

A greater probability of FVF λ_t indicates a larger change in the surrounding space. Therefore, if the value of λ_t is greater than threshold λ_m, the spatial structure is considered to have changed owing to movement. Furthermore, geometrical features $(\mathbf{C}_t^W, \mathbf{S}_t^W)$ in the world coordinate system accumulate as geometrical map $(\mathbf{C}_{t'}^M, \mathbf{S}_{t'}^M)$ passes through voxel-grid filter $f_V(\cdot)$, and the updated time, t', changes to current time t:

$$\mathbf{C}_{t'}^M \leftarrow f_V\left(\mathbf{C}_{t'}^M + \mathbf{C}_t^W\right),\ \mathbf{S}_{t'}^M \leftarrow f_V\left(\mathbf{S}_{t'}^M + \mathbf{S}_t^W\right);\ \lambda_t > \lambda_m, \tag{24}$$

where $(\mathbf{C}_t^W, \mathbf{S}_t^W)$ is calculated using (19) and (20) with $(\mathbf{C}_t^L, \mathbf{S}_t^L)$ and global position $\mathbf{x}_t^W$.

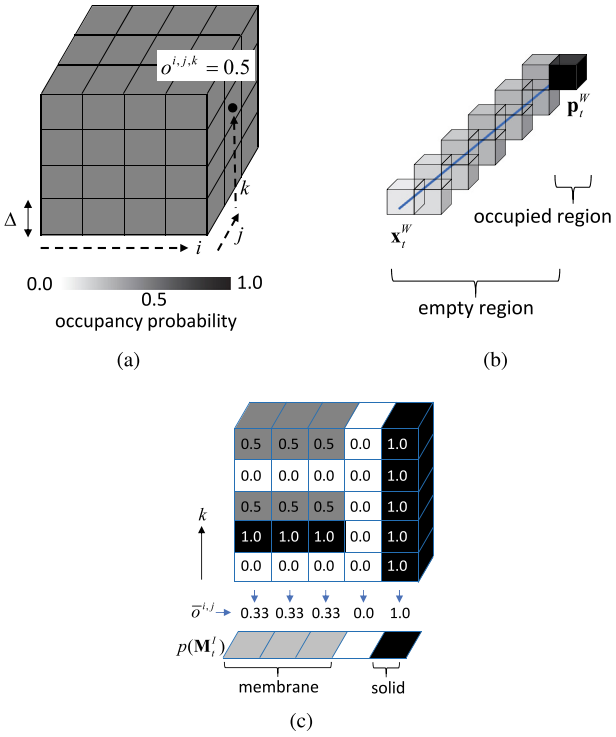


Fig. 4. (a) Voxel map that expresses the presence of an object in a given index, (i, j, k), as an occupancy probability, $o_{i,j,k}$, represents the probability of an object in the cell area between 0 and 1. (b) The LiDAR distance range is divided into an empty region and an occupied region. (c) A 2D plane image with the membrane structure removed can be obtained using the average of the occupancy probability values in the vertical direction.

3.3. Map generation

The global position in the world coordinate system and the 3D points of the node can be used to generate a 3D-point cloud and a 2D-plane image, from which dynamic obstacles are removed and incorporated into the spatial map. The global position and radio signals can also be used to build a radio map. As the constructed spatial and radio maps use the same global position, the coordinate system is automatically linked.

3.3.1. Spatial map

Point cloud $\mathbf{M}_t^P$ accumulated up to time t can be extracted using 3D points $\mathbf{P}_t^L$ of the node and global position $\mathbf{x}_t^W$ as follows:

$$\mathbf{M}_t^P = \mathbf{P}_{1:t}^W,\ \mathbf{P}_t^W = \mathbf{R}(\mathbf{x}_t^W)\left[\mathbf{P}_t^L + \mathbf{T}(\mathbf{x}_t^W)\right],\ p_s(\Delta\mathbf{r}_t) > \lambda_r. \tag{25}$$

By adding 3D points to the point cloud only when spatial probability $p_r(\Delta\mathbf{r}_t)$ of FVF in (23) exceeds threshold λ_r (i.e., the spatial change), the space representation can be compressed using only a few points. In addition, $\mathbf{T}$ and $\mathbf{R}$ are the translation and rotation matrices, respectively, corresponding to global position $\mathbf{x}_t^W$.

The 3D voxel map can be created using 3D points $\mathbf{P}_t^W$ and global position $\mathbf{x}_t^W$, through which the 2D-plane image can be extracted. As shown in Fig. 4(a), the voxel map divides the 3D space into uniform cells of cubic areas and expresses the presence of an object in each cell according to the occupancy probability value. The set of the occupancy probability can be defined as voxel map $\mathbf{M}_t^V$ when 3D points $\mathbf{P}_{1:t}^W$ and global position trajectory $\mathbf{x}_{1:t}^W$ accumulated up to time t are inputted as follows:

$$p(\mathbf{M}_t^V | \mathbf{x}_{1:t}^W, \mathbf{P}_{1:t}^W) = \prod p(o_t^{i,j,k} | \mathbf{P}_{1:t}^W), \tag{26}$$

where index (i, j, k) is determined according to the global position and 3D points. In addition, each cell independently updates the occupancy

probability value regardless of the neighboring cells. Therefore, current occupancy probability $o_t^{i,j,k}$ can be updated using the previous voxel map, $\mathbf{M}_{t-1}^V$, and current 3D point $\mathbf{P}_t^W$ as follows:

$$p(o_t^{i,j,k}|\mathbf{P}_{1:t}^W) = p(o_t^{i,j,k}|\mathbf{P}_t^W, \mathbf{M}_{t-1}^V), \tag{27}$$

where $p(o_t^{i,j,k}|\mathbf{P}_{1:t}^W) = p(o_t^{i,j,k}|\mathbf{P}_t^W)p(o_{t-1}^{i,j,k}|\mathbf{P}_{1:t-1}^W)$ and $p(o_{t-1}^{i,j,k}|\mathbf{P}_{1:t-1}^W) = p(\mathbf{M}_{t-1}^V)$. The ratio of occupied probability $p(o_{t-1}^{i,j,k})$ and emptied probability $p(\bar{o}_{t-1}^{i,j,k})$ in each cell can be defined as an odd value, ξ, which is expressed according to the Bayes rule as follows:

$$\xi = \frac{p(o_t^{i,j,k}|\mathbf{P}_t^W, \mathbf{M}_{t-1}^V)}{p(\bar{o}_t^{i,j,k}|\mathbf{P}_t^W, \mathbf{M}_{t-1}^V)} = p(\mathbf{P}_t^W|o_{t-1}^{i,j,k}) \times \frac{p(o_{t-1}^{i,j,k})}{p(\bar{o}_{t-1}^{i,j,k})}, \tag{28}$$

where $p(\mathbf{P}_t^W|o_{t-1}^{i,j,k})$ is an unknown value and can be defined using an inverse measurement model. This can be designed according to the characteristics of the LiDAR. An inverse measurement model represents the certainty of a measured value, as shown in Fig. 4(b), which depicts the empty and occupied regions, through which the beam passes; these regions are respectively divided into $p_e(\mathbf{P}_t^W|o_{t-1}^{i,j,k})$ and $p_o(\mathbf{P}_t^W|o_{t-1}^{i,j,k})$ and can be calculated as follows:

$$p_e(\mathbf{P}_t^W|o_{t-1}^{i,j,k}) = \frac{d_{i,j,k}}{d_{\max}}, \; p_o(\mathbf{P}_t^W|o_{t-1}^{i,j,k}) = \eta. \tag{29}$$

The empty region is calculated as a value proportional to the maximum detectable distance, $d_{\max}$, which indicates that uncertainty increases with increase in the distance from LiDAR $d_{i,j,k}$. The occupied region is defined as a constant η greater than 1. The greater the value of η, the greater is the change in the occupancy probability. Finally, as the sum of occupied probability $p(o_t^{i,j,k}|\mathbf{P}_t^W, \mathbf{M}_{t-1}^V)$ and emptied probability $p(\bar{o}_t^{i,j,k}|\mathbf{P}_t^W, \mathbf{M}_{t-1}^V)$ becomes 1, the occupancy probability of cell (i, j, k) at time t can be obtained using the odd value as follows:

$$p(o_t^{i,j,k}|\mathbf{P}_t^W, \mathbf{M}_{t-1}^V) = \xi/(1+\xi). \tag{30}$$

Using the generated voxel map, a 2D plane image showing the outline of the space can be extracted. If the occupancy probability value of the voxel map is projected onto the 2D plane in this form, the membrane structures (e.g., ceilings and floors) complicate the clear identification of solid outlines such as walls and columns. To extract a clearer outline, occupancy probability $p(\bar{o}_t^{i,j})$ of the 2D plane image is calculated as the average of the occupancy probability values in the vertical direction of the voxel map as follows:

$$p(\bar{o}_t^{i,j}) = \frac{1}{N_k}\sum_{k=1}^{N_k} p(o_t^{i,j,k}), \; p(o_t^{i,j,k}) \neq 0.5, \tag{31}$$

where the occupancy probability of an unknown area reflects only the occupancy probability of N_k cells instead of 0.5 in the average calculation. Finally, as shown in Fig. 4(c), the 2D plane image, $p(\mathbf{M}_t^I)$, comprises cases in which $p(\bar{o}_t^{i,j})$ exceeds the unknown-area occupancy probability of 0.5:

$$p(\mathbf{M}_t^I) = \prod \left\{ p(\bar{o}_t^{i,j}) \middle| p(\bar{o}_t^{i,j}) > 0.5 \right\}. \tag{32}$$

3.3.2. Radio map

Radio map $\mathbf{M}_t^R$ up to time t used for localization is composed of N_b number of $\mathbf{b}_i$ fingerprints using the radio signal of $\mathbf{Z}_{1:t}$ and global position trajectory $\mathbf{x}_{1:t}^W$ with the same coordinate system as the 2D plane image. This radio map is expressed as

$$\mathbf{M}_t^R = f_r(\mathbf{x}_{1:t}^W, \mathbf{Z}_{1:t}) = \left\{ \mathbf{b}_i \,|i = 1, \ldots, N_b \right\}. \tag{33}$$

The fingerprint is the basic unit of the radio map and contains radio signals $\hat{\mathbf{r}}_i$ of the surrounding Wi-Fi and Bluetooth devices as well as the positions of the key points required for localization in space, and it is defined as follows:

$$\mathbf{b}_i = \left(\hat{x}_i, \hat{y}_i, \hat{\mathbf{r}}_i\right), \hat{\mathbf{r}}_i = \left\{ \left(mc_j, \hat{r}_j\right) | j = 1, \ldots, N_r \right\}. \tag{34}$$

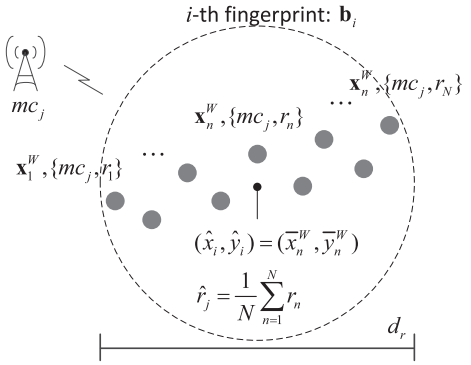


Fig. 5. Generation of a radio-map fingerprint.

In general, RSS values contain unintended errors because of radio signal delays or disturbances caused by spatial obstructions. To reduce the effect of these errors, RSS values of fingerprints are compensated with the average of RSS values obtained from neighboring nodes [40, 41]. In particular, when the placement distances among nearby nodes are large, the location and RSS values of fingerprints are determined by the spatial interpolation modes, such as the Kriging method and Voronoi diagram, to reflect spatial characteristics [41,42]. However, SRS can measure the nodes densely, hence the neighboring nodes include similar spatial information. Therefore, an intermediate location of neighboring nodes can be selected as the fingerprint position. Using the global position, each node can be clustered at a Euclidean distance of d_r by using the k-d tree [38]. As shown in Fig. 5, when N nodes are contained in the ith cluster, fingerprint position $(\hat{x}_i, \hat{y}_i)$ is defined as the middle value of the global position of the nodes in

$$\left(\hat{x}_i, \hat{y}_i\right) = \frac{1}{N}\sum_{n=1}^{N}\left(x_n^W, y_n^W\right), \; \hat{r}_j = \frac{1}{N}\sum_{n=1}^{N} r_n, \tag{35}$$

where RSS value $\hat{r}_j$ corresponding to MAC address mc_j uses the average value of the radio signals. This result is obtained by adopting a median filter to smoothen the noise due to dynamic changes, such as floating population, in the environment.

4. Experimental evaluation

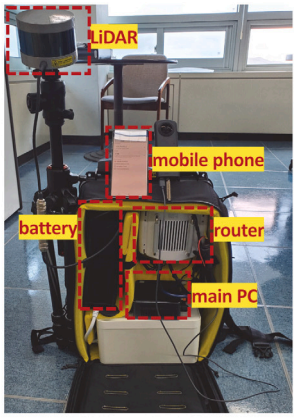
4.1. Experimental setup

To analyze the mapping accuracy, real-time performance, efficiency, and localization applicability as the main contributions of SRS, we constructed a backpack-type mapping system equipped with a main PC, router, mobile phone, and 3D LiDAR, as shown in Fig. 6(a). The backpack was equipped with 16-channel LiDAR that can scan objects at 360° with a speed of 10 Hz. The interval angle between the LiDAR channels was 2° in the vertical direction, and the maximum detectable distance was 100 m. For reference, a typical 3D LiDAR raises the vertical resolution by a multiple of 16 channels according to the internal mirror specification. The higher-resolution channels are more expensive, but LiDAR with over 16 channels basically provides the function to downgrade to 16 vertical resolution in hardware. To assure the backward compatibility of LIDAR, 16-channel LiDAR with a minimum vertical resolution was used for the experiment.

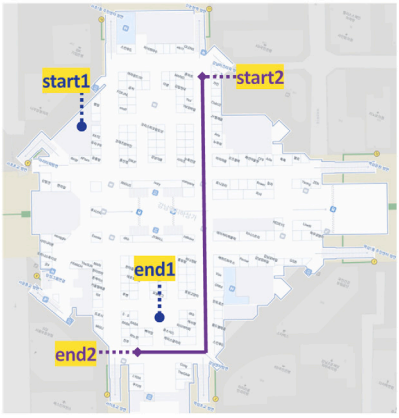
An Android OS-based mobile phone was used that was installed with an app that could collect data on acceleration, angular velocity, magnetic field, atmosphere pressure, and radio signals at 10 Hz. Owing to the slow scan speed of the radio signals, the RSS of the surrounding Wi-Fi and Bluetooth devices was updated at periods of 1 and 5 s, respectively. The main PC was equipped with an Intel i7 3.5-GHz dual-core CPU, with a 16-GB RAM and M.2-type 2-TB SSD.

Table 1
Sensor-specific variance values as the hyperparameters to calculate FVF probability.

	Acc. [m/s^2]	Ang. vel. [rad/s]	Mag. field [μT]	Atm. pres. [hPa]	Radio dist. [m]
Variance	2.32096	0.00085	1.65520	0.00035	705.77690



(a)



(b)

Fig. 6. Experimental setup. (a) Backpack-type portable spatial-tagged radio-mapping system. (b) Departure and arrival points for mapping (start1 and end1); movement trajectory for fingerprinting localization to verify the applicability of the constructed radio map (lines connecting start2 and end2)

The 3D points and radio signals collected from the LiDAR and mobile phone were transmitted to the main PC in real time through LAN and WLAN at 10 Hz. In this case, to resolve the problem of unstable communication of WLAN, TCP/IP was used as the communication protocol. In addition, in the main PC, a new node was created only when the acquisition time of the two data transmitted from the LiDAR and the mobile phone was within 0.1 s. Through this method, even in the case of latency, accurate data synchronization can be achieved.

The experiment was conducted in an indoor shopping center of size 32,040 m^2, as shown in Fig. 6(b). While holding the mobile phone and wearing the backpack, the operator moved 1.75 km from the designated starting point to the destination point and acquired 15,301 nodes. The nodes are used for generating a 3D-point cloud, 2D-plane image, and radio map linked to the unified coordinate system, as shown in Algorithm 1. In addition, the number of nearby Wi-Fi routers in the space affects the fingerprinting localization [43]. When the number of Wi-Fi routers is less than three, fingerprinting localization becomes severely inaccurate, and Wi-Fi routers with over 100 cause the sensing delay phenomenon [44,45]. This experimental site is suitable for evaluating radio mapping and fingerprinting localization in terms of accuracy and real-time performance because 47 Wi-Fi signals were detected per node on average.

Table 2
Accuracy of a spatial map for each section (unit: [m]).

Section	Map distance	Real distance	Error
A	22.25	22.20	0.05
B	43.75	43.72	0.03
C	48.90	48.96	−0.06
D	51.20	51.18	0.02
E	36.45	39.52	−0.07
F	48.70	48.73	−0.03
G	121.75	121.8	−0.05

Table 1 shows the variance of the probability density function used for each sensor to calculate the probability of the FVF to improve the mapping efficiency. In Eq. (22), the variances are the hyperparameters that need to be predetermined before operating the FVF. To ensure generality, the variances of measurements obtained from sensors on mobile phones were used. Actually, the motion data and radio signals obtained in 10 spaces, including offices, shopping centers, and subway stations, were analyzed and calculated the sensor-specific variance values for FVF probability. In addition, the threshold of the FVF probability was set at 0.65 for detecting motion and spatial changes in the surrounding area. The larger the threshold value, the larger the spatial change required for the FVF to store geometrical features in the geometrical map.

4.2. Experimental results

To analyze the performance of the SRS, an experiment was conducted considering three cases: Epoch1, Epoch3, and FVF. In Epoch1, all the generated geometrical features were accumulated in the geometrical map and used. In Epoch3, only one of three nodes were accumulated, and the starting and destination points for the mapping were displayed as circular points, as shown in Fig. 6(b).

4.2.1. Mapping accuracy

Fig. 7 shows the 3D-point cloud and 2D-plane image generated through the FVF. As shown, the generated 2D-plane image and outline of the actual experimental space in Fig. 6(b) are similar. To evaluate the accuracy of the 2D-plane images, the real distance for each section in Fig. 7(b) was measured using plus total station. Consequently, as shown in Table 2, we can see that an error of up to 7 cm occurs in the map accuracy depending on the section. In addition, to analyze the SRS radio map accuracy, Fig. 8(a) shows the X–Y axis displaying the 2D movement trajectory of the three modes. In Fig. 8(a), Epoch1 and FVF appear as one movement trajectory with nearly identical localization performances for mapping. In contrast, in the case of Epoch3, localization failed during mapping execution. To reduce the processing time, Epoch3 updated the geometrical map according to a fixed period; the key geometrical features of the process were not included in the map. Consequently, an erroneous matching occurred. Conversely, FVF probabilistically analyzed the operator's motion and changes in the surrounding space and uniformly registered geometrical features for all areas in the map; this enabled the FVF to achieve localization accuracy

Algorithm 1 Mapping Process of Spatial-tagged Radio Mapping System

Input: 3D point $\mathbf{P}_t$ and radio signal $\mathbf{Z}_t$
Output: 3D-point cloud $\mathbf{M}_t^P$, 3D voxel map $\mathbf{M}_t^V$,2D plane image $\mathbf{M}_t^I$, radio map $\mathbf{M}_t^R$
Parameter: threshold of feature variance filter probability λ_m

begin
  for new input data $t \leftarrow t+1$ **do**
    <**Synchronization**>
    node generation: $\mathbf{N}_t = \{\mathbf{P}_t^L, \mathbf{Z}_t\}$
    <**Localization for Map Construction**>
    local point and geometrical features extraction: $(\mathbf{V}_t^L, \mathbf{H}_t^L)$, $(\mathbf{C}_t^L, \mathbf{S}_t^L)$
    local position estimation through scan matching: $\mathbf{x}_t^L$
    global position prediction: $\tilde{\mathbf{x}}_t^W$
    point features transformation to world: $(\mathbf{V}_t^W, \mathbf{H}_t^W)$
    if feature variance filter: $\lambda_t > \lambda_m$ **then**
      geometrical features transformation to world: $(\mathbf{C}_t^W, \mathbf{S}_t^W)$
      geometrical map update using voxel grid filter: $\mathbf{C}_{t'}^M \leftarrow f_V\left(\mathbf{C}_{t'}^M + \mathbf{C}_t^W\right)$, $\mathbf{S}_{t'}^M \leftarrow f_V\left(\mathbf{S}_{t'}^M + \mathbf{S}_t^W\right)$
      time update of geometrical map: $t' \leftarrow t$
    end if
    global position estimation through map matching: $\mathbf{x}_t^W$
    <**Map Generation**>
    point cloud: $\mathbf{M}_t^P \leftarrow \mathbf{M}_{t-1}^P + \mathbf{P}_t^W$
    voxel map using point cloud: $\mathbf{M}_t^V = \prod p(o_t^{i,j,k} | \mathbf{P}_t^W, \mathbf{M}_{t-1}^V)$
    plane image using occupancy probability of voxel map: $p(\mathbf{M}_t^I) = \prod p(\bar{o}_t^{i,j})$; $p(\bar{o}_t^{i,j}) = \sum_{k=1}^{N_k} p(o_t^{i,j,k})/N_k$, $p(o_t^{i,j,k}) \neq 0.5$
    radio map using radio signal: $\mathbf{M}_t^R = f_r(\mathbf{x}_{1:t}^W, \mathbf{Z}_{1:t})$
  end for
end

Table 3
Position errors of Epoch3 and FVF compared to those of Epoch1 for mapping after moving 1.75 km (unit: [m] or [rad]).

Moving distance	Mode	X	Y	Z	Yaw	Pitch	Roll
500	Epoch3	3.75	−8.19	−0.12	0.05	−0.04	−0.05
	FVF	0.05	−0.09	−0.03	0	0	0.01
1000	Epoch3	−19.54	15.28	−2.59	0.09	0.02	−0.51
	FVF	0.07	0.06	−0.06	−0.01	0.01	0
1500	Epoch3	15.45	41.38	−1.45	0.07	0.05	−0.95
	FVF	**0.12**	0.05	−0.04	**0.03**	0.02	0.02
1750	Epoch3	27.3	**56.69**	−0.98	0.09	0.08	**−5.67**
	FVF	0.1	0.08	−0.01	0.02	0	−0.01

for mapping similar to that of Epoch1. Table 3 lists the numerical analysis results of the 6-DoF localization accuracy for mapping on the movement trajectory. Based on Epoch1, the results show the difference between the estimated positions of Epoch3 and FVF for each 500 m when operator moved 1.75 km. As observed, FVF showed errors within 0.12 m and 0.03 rad, while Epoch3 diverged up to 56.69 m and 5.67 rad.

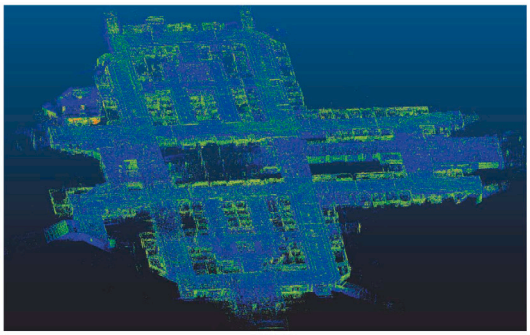
4.2.2. Real-time performance

Real-time evaluation was performed by measuring the mapping speed while verifying the mapping accuracy; this is described in Section 4.2.1 using the main PC described in Section 4.1. Considering that the main factors affecting the computation speed of the three methods such as Epoch1, Epoch3, and FVF are the number of features in the geometrical map, an experiment was conducted to analyze the relationship between the number of registered features in the geometrical map and the processing time.

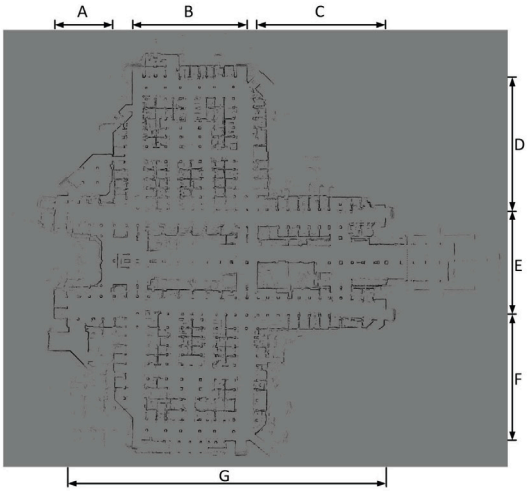
As shown in Fig. 9(a), feature points and geometrical features can be generated in the local coordinate system according to the input points obtained from LiDAR. The results demonstrated that the point and geometrical features are uniformly extracted regardless of the operator's movement and changes in the surrounding space. In the local coordinate system, the LiDAR measured an average of 24,491 points per node as mapping input data. As the feature points, the corner and the planar points were generated by 1194 and 4697, respectively. As the geometrical features for scan matching, the corner lines and the planar surfaces were generated by 596 and 1565, respectively. In the world coordinate system, as shown in Fig. 9(b), the number of features entered in the geometrical map according to Epoch1, Epoch3, and FVF differ. In Epoch1, all the generated geometrical features are registered in the map, whereas in Epoch3, they are registered with a period of 1/3 nodes to reduce the computational burden. In Epoch1 and Epoch3, geometrical features of 15,301 and 5100 nodes, respectively, were stored in the geometrical map for map matching. In FVF, only the features of the geometrical map with 3772 nodes are updated; this is the point at which the moving space is determined to have changed, as represented by the cross marks in Fig. 8(b).

Comparing Epoch1 and Epoch3 in Fig. 9(b), the number of the geometrical features is not linearly proportional to thrice the difference, which is the number of nodes used in building the geometrical map, because of the voxel grid filter common to all methods. Specifically, the leaf sizes of the voxel-grid filter for the corner lines and planar surfaces were 0.2 and 0.4 m, respectively. The voxel grid filter prevents locally overlapping features from registering them in the geometrical map. In other words, there is a limit to effectively reducing the number of features in the geometrical map only with a simple periodic method, such as Epoch3. The number of features in the geometrical map sharply increased with Epoch1, followed by Epoch3, and finally FVF, confirming the effect of FVF.

Fig. 10 shows the processing times required for one node of the actual mapping with 6-DoF localization. The average processing times in Epoch1 and Epoch3 were 2.8813 and 11.4298 Hz, respectively. Comparatively, FVF was faster at 19.2432 Hz. This is because of two major



(a)



(b)

X	Y	Radio #	MAC1	RSS1	MAC2	RSS2	MAC3	RSS3	
0.465	-0.265	61	c4ad3421a88d	-44	c4ad3421a88c	-45	0009b4752923	-56	...
1.445	-0.47	66	c4ad3421a88d	-38	c4ad3421a88c	-44	0009b4752923	-56	
2.365	-0.105	45	c4ad3421a88d	-38	c4ad3421a88c	-42	0009b4752923	-56	
...									

(c)

Fig. 7. Mapping result of the (a) 3D-point cloud, (b) 2D-plane image, and (c) radio map.

factors: (1) the voxel grid filter requires classifying the features into the regions when updating the geometrical map, (2) the k-d tree algorithm needs to search the correspondences between the current observation features and the geometrical map when performing the map matching. The voxel grid filter and k-d tree algorithm are affected by the number of nodes used for composing with the geometrical map and the number of features in the geometrical map, respectively. Therefore, the FVF with the smallest number of nodes and features composing the geometrical map shows the best performance in processing time. It can be experimentally confirmed that the FVF offers a significant contribution in terms of real-time performance. Considering that the scan speed of LiDAR is 10 Hz, Epoch3 secured real-time performance; however, the accuracy estimation confirmed the possibility of erroneous localization. Therefore, it can be said that only the FVF operates in real time with an accuracy similar to that of Epoch1.

4.2.3. Mapping efficiency

For experimental validation, radio maps were constructed using SRS and PMC, as shown in Fig. 7(c). According to the results, SRS requires 50 min to extract 984 fingerprints at 1-m intervals on the 2D-plane image from the collected node data, as shown in Fig. 8(b). This time includes the operation conducted to generate the 2D-plane image. This result could be achieved because the spatial probability of FVF allows only the point cloud data of 950 nodes in space (Fig. 8(b)) to be utilized for building the 2D-plane image. In contrast, PMC required 4 h owing to manual work. Arithmetically, SRS exhibited a mapping efficiency 4.8 times faster than that of PMC. These experimental results demonstrate that SRS can not only generate 2D-plane image but also achieve better mapping efficiency than PMC.

4.2.4. Applicability to LBS

To verify the SRS applicability of the constructed radio map with respect to the LBS, an experiment was conducted to analyze the accuracy of fingerprinting localization and check whether heterogeneous mobile phones are supported. Even in the case of indoor localization competitions, high cost is required because special equipment must be used to produce reference datasets containing ground-truth information [46,47]. Therefore, in this study, as can be seen in Fig. 6(b), a diamond-shaped path consisting of straight line sections was accurately calibrated and adjusted using the manual method of PMC to construct reference datasets. Additionally, for the analyzing the accuracy of localization, the empirical-model-based Radar [48] and statistical-model-based Horus [49], which are widely used in the LBS application field, were adopted as fingerprinting localization methods.

Fig. 11(a) shows the localization accuracies when using the radio maps constructed by SRS and PMC. Objectively, to compare the applicability of maps generated by SRS and PMC, two radio maps were

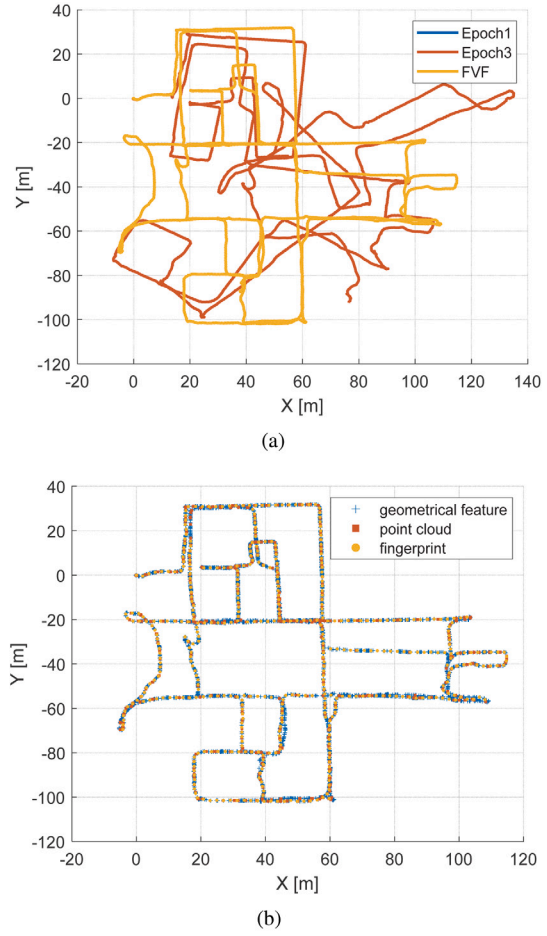


Fig. 8. (a) Localization trajectory for mapping. (b) Major nodes for mapping by FVF: 3772 nodes for geometrical features, 950 nodes for point cloud data, 984 nodes for fingerprints at 1-m intervals.

converted with the same procedure according to the naïve procedure of Radar and Horus. Both Radar and Horus methods employ the tracking effects of radio signals, and they compute the mean RSS values as the observation using the sequential radio signals. The RSS observation is used for matching with the radio map to perform the LBS. In the experiment, Radar and Horus utilized 10 and 2 radio signals, respectively, in the past to calculate the RSS observation that were equally applied to all fingerprinting localization. The reason for using the same tracking parameters was to verify the LBS applicability of radio maps constructed by SRS compared with PMC, rather than to improve the fingerprinting localization algorithm.

The SRS radio map did not result in a large difference in localization performance compared to that of the PMC radio map. In the case of the Radar method, the use of SRS and PMC achieved an error of 3.83491 and 4.18096 m, respectively, at the 80th percentile; therefore, the SRS showed a slightly higher accuracy. In the case of combination with inertial-based PDR, more accurate localization results can be obtained, but there is a limitation in analyzing the effect of pure SRS radio maps that affect the localization accuracy. Considering that the 75th percentile mean accuracy of research teams using the inertial-based PDR in combination with fingerprinting algorithms in the recent well-known LBS competition is 5.4 m, it is believed that the applicability of the proposed SRS radio map to the actual LBS is high [30]. Fig. 11(b) shows the localization results obtained with application of the SRS-based radio map to other types of mobile phones. The trend of the localization accuracy was divided into two categories according to the Radar and Horus algorithms. However, the localization performance did not significantly differ according to the device. Thus, the experimental results verify the practical applicability of the SRS-generated radio maps in terms of localization accuracy and heterogeneous mobile phone support.

5. Conclusions and discussions

In this study, we proposed an SRS that can simultaneously generate spatial and radio maps even in large indoor environments. The SRS effectively fuses the sensing data obtained from two devices, i.e., a LiDAR scanner and a mobile phone, to generate spatial and radio maps in real time. The novelty of SRS is that it complementarily uses the distance range of LiDAR and radio signals and motion data of the mobile phone to secure high mapping accuracy and real-time performance. SRS was created in the form of a portable backpack, and experiments were performed to generate maps on the field. The results demonstrated that SRS achieves high accuracy, real-time performance, efficiency advancement, and localization applicability in both spatial and radio mapping. Furthermore, the applicability of the SRS to actual LBS was confirmed by utilizing the constructed maps to well-known fingerprinting localization algorithms and heterogeneous mobile phones.

SRS requires specialized equipment such as LiDAR, thus it is suitable for creating an initial radio map in an unknown environment by map building operators instead of public users. Considering the accessibility of SRS equipment, in the case of a site where the radio maps are provided in advance, the cloud collection method that can update the existing maps with only mobile phones might be more effective than SRS in terms of map management cost.

In the future works, therefore, we will aim to add a function to update the radio map constructed by SRS using a cloud strategy. Because the cloud method constructs radio maps in various ways, uniform accuracy cannot be ensured. Nevertheless, with the cloud method, if high-accuracy radio maps are available, the maps can be updated by continuously adapting to spatial changes. Accordingly, we need to adopt a strategy in which SRS is used for providing initial radio map with high accuracy and cloud method is used for updating the changed radio map. In terms of the continuity of the radio map constructed by SRS, the use of a cloud strategy in updating radio map is expected to serve as a highly effective method for practical indoor LBS.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Yu-Cheol Lee reports financial support was provided by Korea Ministry of Science and ICT.

Acknowledgments

We would like to express our deepest thanks to Peter KS Kim and Kibeom Kim for their assistance in providing useful comments as the reader and collecting the dataset. We are also grateful to the reviewers, whose comments and suggestions have helped to improve and clarify this manuscript.

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Ministry of Science and ICT of South Korea (No. 2018-0-00205, Development of Core Technology of Robot Task-Intelligence for Improvement of Labor Condition).

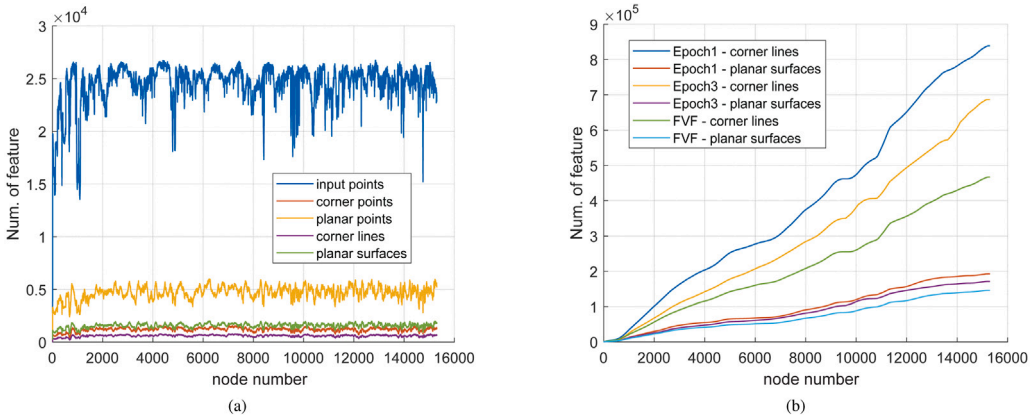


Fig. 9. (a) Number of input points, point features, and geometrical features in the local coordinate system; (b) number of features in the geometrical map in the world coordinate system.

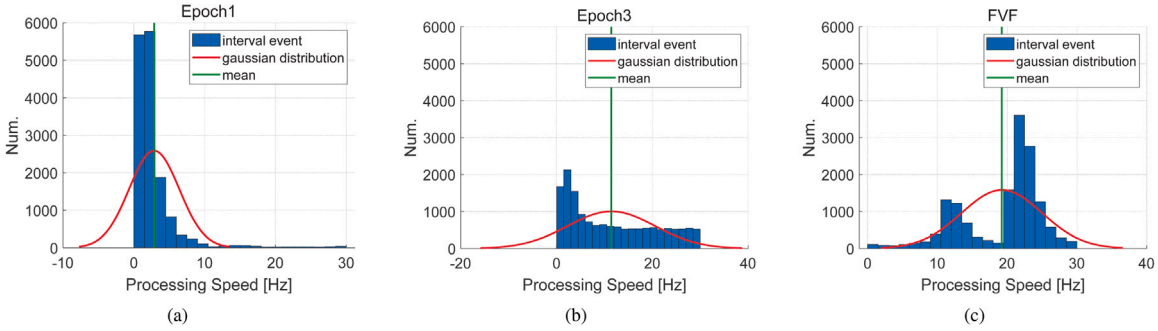


Fig. 10. Analysis of processing time required for one node of mapping: (a) Epoch1 (mean: 2.8813, variance: 12.7707), (b) Epoch3 (mean: 11.4298, variance: 88.2202), (c) FVF (mean: 19.2432, variance: 33.4038)

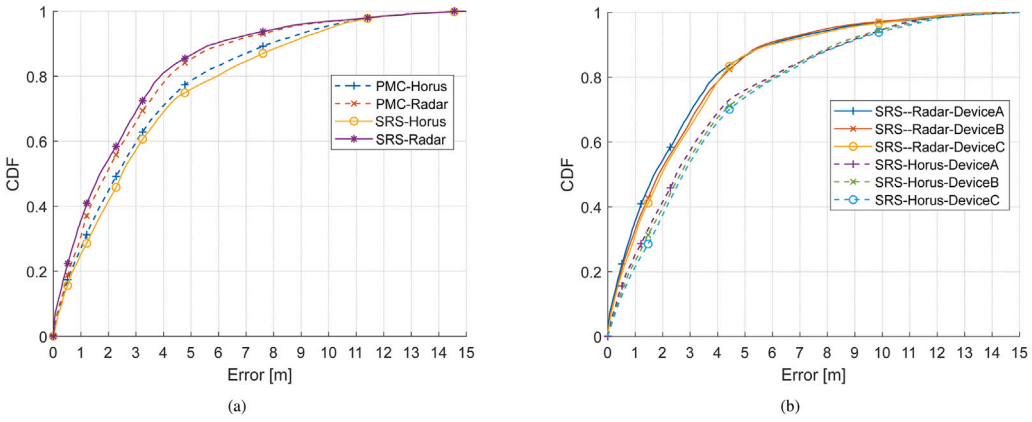


Fig. 11. Localization results using fingerprinting algorithms, Horus and Radar. (a) Location accuracy using radio maps constructed by PMC and SRS. (b) Location accuracy using the SRS radio map and a heterogeneous device.

References

[1] X. Wang, X. Wei, Y. Liu, S. Gao, Received signal strength–based localization for large space indoor environments, Int. J. Distrib. Sens. Netw. 13 (1) (2017) 1–12.

[2] K.W. Kolodziej, J. Hjelm, Local Positioning Systems: LBS Applications and Services, CRC Press, 2017.

[3] A.F.G. Ferreira, D.M.A. Fernandes, A.P. Catarino, J.L. Monteiro, Localization and positioning systems for emergency responders: A survey, IEEE Commun. Surv. Tutor. 19 (4) (2017) 2836–2870.

[4] W. Sun, M. Xue, H. Yu, H. Tang, A. Lin, Augmentation of fingerprints for indoor WiFi localization based on Gaussian process regression, IEEE Trans. Veh. Technol. 67 (11) (2018) 10896–10905.

[5] A. Achroufene, Y. Amirat, A. Chibani, RSS-Based indoor localization using belief function theory, IEEE Trans. Autom. Sci. Eng. 16 (3) (2018) 1163–1180.

[6] M. Kolakowski, V. Djaja-Josko, TDOA-TWR Based positioning algorithm for UWB localization system, in: Proc. of IEEE MIKON, Krakow, Poland, 2016, pp. 1–4.

[7] J. Xu, M. Ma, C.L. Law, Cooperative angle-of-arrival position localization, Measurement 59 (2015) 302–313.

[8] Z. Xu, Z. Liang, Z. Zhou, Z. Li, G.-M. Du, X. Wang, Y.-K. Song, A precise 3D positioning approach based on UWB with reduced base stations, in: Proc. of IEEE ASID, Xiamen, China, 2021, pp. 145–149.

[9] R. Bazo, C.A. da Costa, L.A. Seewald, L.G. da Silveira, R.S. Antunes, R. da Rosa Righi, V.F. Rodrigues, A survey about real-time location systems in healthcare environments, J. Med. Syst. 45 (3) (2021) 1–13.

[10] C. Wu, X. Wang, M. Chen, M.J. Kim, Differential received signal strength based RFID positioning for construction equipment tracking, Adv. Eng. Inf. 42 (2019) 100960.

[11] S. Li, Z. Qin, H. Song, A temporal-spatial method for group detection, locating and tracking, IEEE Access 4 (2016) 4484–4494.

[12] H. Zou, M. Jin, H. Jiang, L. Xie, C.J. Spanos, WinIPS: Wifi-based non-intrusive indoor positioning system with online radio map construction and adaptation, IEEE Trans. Wirel. Commun. 16 (12) (2017) 8118–8130.
[13] Y. Shu, et al., Gradient-based fingerprinting for indoor localization and tracking, IEEE Trans. Ind. Electron. 63 (4) (2016) 2424–2433.
[14] H.-H. Liu, The quick radio fingerprint collection method for a wifi-based indoor positioning system, Mob. Netw. Appl. 22 (1) (2017) 61–71.
[15] Y.-C. Lee, H. Myung, Indoor localization method based on sequential motion tracking using topological path map, IEEE Access 7 (2019) 46187–46197.
[16] C. Gao, R. Harle, Easing the survey burden: Quantitative assessment of low-cost signal surveys for indoor positioning, in: Proc. of IEEE IPIN, Alcala de Henares, Spain, 2016, pp. 1–8.
[17] X. Luo, W.J. O'Brien, C.L. Julien, Comparative evaluation of received signal-strength index (RSSI) based indoor localization techniques for construction jobsites, Adv. Eng. Inf. 25 (2011) 355–363.
[18] Z. Yang, C. Wu, Y. Liu, Locating in fingerprint space: wireless indoor localization with little human intervention, in: Proc. of ACM MOBICOM, Istanbul, Turkey, 2012, pp. 269–280.
[19] Y. Li, et al., A hybrid WiFi/magnetic matching/PDR approach for indoor navigation with smartphone sensors, IEEE Commun. Lett. 20 (1) (2016) 169–172.
[20] S.H. Jung, B.-C. Moon, D. Han, Performance evaluation of radio map construction methods for Wi-Fi positioning systems, IEEE Trans. Intell. Transp. Syst. 18 (4) (2016) 880–889.
[21] L. Bruno, P. Robertson, WiSLAM: Improving footslam with wifi, in: Proc. of IEEE IPIN, Guimaraes, Portugal, 2011, pp. 1–10.
[22] J.J. Pan, et al., Tracking mobile users in wireless networks via semi-supervised colocalization, IEEE Trans. Pattern Anal. Mach. Intell. 34 (3) (2011) 587–600.
[23] K. Chintalapudi, A. Padmanabha Iyer, V.N. Padmanabhan, Indoor localization without the pain, in: Proc. of ACM MobiCom, Chicago IL, USA, 2010, pp. 173–184.
[24] I. Bisio, F. Lavagetto, M. Marchese, A. Sciarrone, GPS/HPS-and Wi-Fi Fingerprint-based location recognition for check-in applications over smartphones in cloud-based LBSs, IEEE Trans. Multimed. 15 (4) (2013) 858–869.
[25] J. Yu, Z. Na, X. Liu, Z. Deng, WiFi/PDR-integrated indoor localization using unconstrained smartphones, EURASIP J. Wireless Commun. Networking 2019 (1) (2019) 1–13.
[26] J. Kunhoth, A. Karkar, S. Al-Maadeed, A. Al-Ali, Indoor positioning and wayfinding systems: a survey, Hum.-Centric Comput. Inf. Sci. 10 (2020) 1–41.
[27] H. Ju, S.Y. Park, C.G. Park, A smartphone-based pedestrian dead reckoning system with multiple virtual tracking for indoor navigation, IEEE Sens. J. 18 (16) (2018) 6756–6764.
[28] H.-C. Yen, C.-C. Wang, C.-F. Chou, Orientation constraints for Wi-Fi SLAM using signal strength gradients, Auton. Robot. 44 (2020) 1–12.
[29] R. Liu, et al., Collaborative SLAM based on WiFi fingerprint similarity and motion information, IEEE Internet Things J. 7 (3) (2019) 1826–1840.
[30] V. Renaudin, et al., Evaluating indoor positioning systems in a shopping mall: The lessons learned from the IPIN 2018 competition, IEEE Access 7 (2019) 148594–148628.
[31] H. Liu, et al., Accurate WiFi based localization for smartphones using peer assistance, IEEE Trans. Mob. Comput. 13 (10) (2014) 2199–2214.
[32] D. Denkovski, V. Atanasovski, L. Gavrilovska, J. Riihijärvi, P. Mähönen, Reliability of a radio environment map: Case of spatial interpolation techniques, in: Proc. of IEEE CROWNCOM, Stockholm, Sweden, 2012, pp. 248–253.
[33] S. Sorour, Y. Lostanlen, S. Valaee, K. Majeed, Joint indoor localization and radio map construction with limited deployment load, IEEE Trans. Mob. Comput. 14 (5) (2015) 1031–1043.
[34] J.S. Dai, Euler–rodrigues formula variations, quaternion conjugation and intrinsic connections, IEEE Trans. Mob. Comput. 92 (2015) 144–152.
[35] K. Levenberg, A method for the solution of certain non-linear problems in least squares, Q. Appl. Math. 2 (2) (1944) 164–168.
[36] J.J. Moré, The levenberg-marquardt algorithm: implementation and theory, in: Numerical Analysis, Dundee, UK, 1978, pp. 105–116.
[37] M. Magnusson, N. Vaskevicius, T. Stoyanov, K. Pathak, A. Birk, Beyond points: Evaluating recent 3D scan-matching algorithms, in: Proc. of IEEE ICRA, Seattle, WA, USA, 2015, pp. 3631–3637.
[38] M. Otair, Approximate k-nearest neighbour based spatial clustering using kd tree, Int. J. Database Manag. Syst. 5 (1) (2013) 97–108.
[39] M. Shchekotov, Indoor localization method based on Wi-Fi trilateration technique, in: Proc. of FRUCT, St.-Petersburg, Russia, 2014, pp. 177–179.
[40] W. Xue, W. Qiu, X. Hua, K. Yu, Improved Wi-Fi RSSI measurement for indoor localization, IEEE Sens. J. 17 (7) (2017) 2224–2230.
[41] J. Zuo, S. Liu, H. Xia, Y. Qiao, Multi-phase fingerprint map based on interpolation for indoor localization using iBeacons, IEEE Sens. J. 18 (8) (2018) 3351–3359.
[42] N. Swangmuang, P. Krishnamurthy, An effective location fingerprint model for wireless indoor localization, Pervasive Mob. Comput. 4 (6) (2008) 836–850.
[43] Y. Tian, B. Huang, B. Jia, L. Zhao, Optimizing AP and beacon placement in WiFi and BLE hybrid localization, J. Netw. Comput. 164 (2020) 102673.
[44] B. Jia, B. Huang, H. Gao, W. Li, L. Hao, Selecting critical WiFi APs for indoor localization based on a theoretical error analysis, IEEE Access 7 (2019) 36312–36321.
[45] J. Jun, L. He, Y. Gu, W. Jiang, G. Kushwaha, A. Vipin, L. Cheng, C. Liu, T. Zhu, Low-overhead WiFi fingerprinting, IEEE Trans. Mob. Comput. 17 (3) (2017) 590–603.
[46] M. Angermann, P. Robertson, T. Kemptner, M. Khider, A high precision reference data set for pedestrian navigation using foot-mounted inertial sensors, in: Proc. of IEEE IPIN, Zurich, Switzerland, 2010, pp. 1–6.
[47] D. Lymberopoulos, J. Liu, The microsoft indoor localization competition: Experiences and lessons learned, IEEE Signal Process Mag. 34 (5) (2017) 125–140.
[48] P. Bahl, V.N. Padmanabhan, RADAR: An in-building RF-based user location and tracking system, in: Proc. of IEEE INFOCOM, Tel Aviv, Israel, 2000, pp. 775–784.
[49] M. Youssef, A. Agrawala, The horus WLAN location determination system, in: Proc. of ACM MOBISYS, Seattle WA, USA, 2005, pp. 205–218.